MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 3, 2021/2022

DSP5614 – PROBABILITY AND STATISTICS

(*for diploma students only*)

10 AUGUST 2022

9.00 a.m. – 11.00 a.m.
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of **3** pages excluding cover page and appendix.
2. Attempt **ALL FOUR (4)** questions. The marks distribution for each question is given.
3. Write your answers in the answer booklet provided.
4. Key formulae and statistical tables are given in the Appendix.

Question 1

a) Given a data set of time taken by 30 random beach goers to swim 50m.

65	73	59	58	33	70	56	71	55	73
48	52	67	46	75	51	75	55	32	54
76	78	34	62	68	79	39	72	64	46

i) Construct a grouped data table starting with 30⩽40. (4 marks)

ii) Based on the grouped data table in (i), calculate the sample mean and standard deviation. (7 marks)

b) The following stem-and-leaf diagram shows marks of 20 students in a statistics class.

STEM	LEAF
3	4 9
4	3
5	4
6	7
7	0 5 6
8	1 2 4 5 6 8 9 9
9	0 6 6 9

Key: 7|0 means 70

i) Identify the five number summary. (8 marks)

ii) Calculate the lower and upper limit of the distribution. (3 marks)

iii) Identify any outliers and draw a boxplot to represent the above data. (4 marks)

iv) Calculate the sample mean. (2 marks)

v) Determine the shape of the distribution based on answers in (i) and (iv). (2 marks)

[TOTAL 30 MARKS]

Continued...

Question 2

a) A research done by marketing department of a smartphone company have identified the failure rate of their product by market condition. The market condition consists of Favourable, Neutral and Unfavourable where the percentage of the condition are 47%, 33%, and 20% respectively. The failure rate turns out to be 1%, 2%, and 4% for those three market condition, respectively.

i) Construct a tree diagram to represent the above information. (4 marks)

ii) What is the probability of a successful product? (4 marks)

iii) If the product is successful, what is the probability that the market condition was Favorable? (3 marks)

b) A survey has been carried out among 750 undergraduates in a Creative Media University for their participation in a video making competition. The survey result showed that 525 undergraduates have participated in a short story competition and 475 undergraduates participated in commercial competition. However, there are 325 undergraduates that have participated in both short story and commercial competition. If an undergraduate is chosen at random, what is the probability that this undergraduate has participated in either a short story or a commercial competition. (4 marks)

c) The table below shows the registration for marathon and half-marathon event according to age group.

	Half-Marathon	**Full Marathon**
Veteran (V)	47	85
Adults (A)	102	98
Youth (Y)	101	67

A participant is selected randomly after the event. Calculate the probability that the selected participant

i) is an adult. (2 mark)

ii) registered for half-marathon, given that the participant is a youth. (3 marks)

[TOTAL 20 MARKS]

Question 3

a) A normal coffee plant will produce an average of 8.3kg of coffee beans with a standard deviation of 2.4kg. A plant scientist claimed that more coffee can be produced by using enhanced crops. To check on this, a random sample of 36 plants were tested and their mean produce is 9.1kg. At 1% significance value, can we accept the claim by the plant scientist? (10 marks)

Continued...

b) A survey conducted by Human Resource Department was trying to validate that more than 70% of their married staff, has children. The number of children of their married staff were recorded as below:

0	0	1	1	2	0	2	6	0	4
2	2	0	2	0	1	2	2	3	0
0	1	2	5	1	1	3	3	0	4
1	2	2	2	3	0	0	3	3	1

Based on the result above, can we conclude at 10% significance level, that more than 70% of their married staff, has children? (12 marks)

c) If a random sample of 35 subjects in a two tailed hypothesis test, results in a test statistics value of $z_0 = -1.75$. Calculate the p-value. (3 marks)

[TOTAL 25 MARKS]

Question 4

a) A student in Mr Nabil's Statistics class is studying the relationship between height and weight of large shipment package. Below are the data collected.

Height (mm), x	Weight (g), y
63	127
64	121
66	142
69	157
69	162
71	156
71	169
72	165
73	181
75	208

i) Calculate SS_{xx}, SS_{yy} and SS_{xy} based on the information given. (9 marks)

ii) Calculate the correlation coefficient, r and interpret the answer. (4 marks)

iii) Find the linear regression equation $\hat{y} = \beta_0 + \beta_1 x$. (5 marks)

iv) Predict the weight of the shipment package if the height is 70mm. (2 marks)

v) Compute the value of *SST*, *SSR* and *SSE*. (3 marks)

vi) Calculate the standard error of the estimate, s_e. (2 marks)

[TOTAL 25 MARKS]

End of page.

APPENDIX I

Key Formula

	Ungrouped Data	Grouped Data
Mean	$\bar{x} = \frac{\sum x_i}{n}$ where n : sample size	$\bar{x} = \frac{\sum m_i f_i}{\sum f_i}$ where m : class midpoint f : class frequency
Variance	$s^2 = \frac{1}{n-1}\left[\sum x_i^{\,2} - \frac{\left(\sum x_i\right)^2}{n}\right]$ where n : sample size	$s^2 = \frac{1}{\left(\sum f_i\right)-1}\left[\sum m_i^{\,2} f_i - \frac{\left(\sum m_i f_i\right)^2}{\sum f_i}\right]$ where m : class midpoint f : class frequency

Standardized Normal Distribution

- z-value (observed value) for an x value : $Z = \frac{x-\mu}{\sigma}$

Hypothesis Testing

Population	Mean	Proportion
1	$Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ where $\hat{p} = \frac{x}{n}$
2	$Z = \frac{\left(\bar{x}_1-\bar{x}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^{\,2}}{n_1}+\frac{\sigma_2^{\,2}}{n_2}}}$	$Z = \frac{\left(\hat{p}_1-\hat{p}_2\right)-\left(p_1-p_2\right)}{\sqrt{\hat{p}_p\left(1-\hat{p}_p\right)}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ *where* $\hat{p}_p = \frac{x_1+x_2}{n_1+n_2}$

Continued...

Linear Regression Analysis

- **Least square regression equation,** $\hat{y} = \beta_0 + \beta_1 x$ where $\beta_1 = \dfrac{S_{xy}}{S_{xx}}$ and $\beta_0 = \bar{y} - \beta_1 \bar{x}$

- **Correlation coefficient,** $r = \dfrac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$ where

$S_{xx} = \sum x^2 - \dfrac{(\sum x)^2}{n}$	$S_{yy} = \sum y^2 - \dfrac{(\sum y)^2}{n}$	$S_{xy} = \sum xy - \dfrac{(\sum x)(\sum y)}{n}$

- **Sum of Square**

$SST = SSR + SSE$	$SST = S_{yy}$	$SSR = \dfrac{(S_{xy})^2}{S_{xx}}$

- **Test of Significance for Regression Slope**

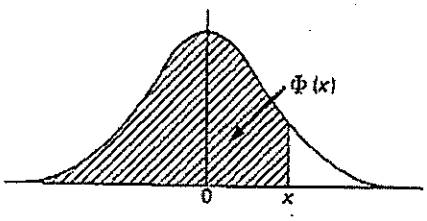
Standard Error, $s_e = \sqrt{\dfrac{SSE}{n-2}}$	Regression t-Test $t = \dfrac{\beta_1}{s_e / \sqrt{Sxx}}$	Correlation t-Test $t = r \cdot \sqrt{\dfrac{n-2}{1-r^2}}$

Continued...

APPENDIX II

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.



x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158

Continued...

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

End of page.